



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Prefixal agreement and impersonal 'il' in Spoken French: Experimental evidence

**Citation for published version:**

Culbertson, J & Legendre, G 2014, 'Prefixal agreement and impersonal 'il' in Spoken French: Experimental evidence', *Journal of French Language Studies*, vol. 24, no. 01, pp. 83-105.  
<https://doi.org/10.1017/S0959269513000380>

**Digital Object Identifier (DOI):**

[10.1017/S0959269513000380](https://doi.org/10.1017/S0959269513000380)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Journal of French Language Studies

**Publisher Rights Statement:**

© Culbertson, J., & Legendre, G. (2014). Prefixal agreement and impersonal 'il' in Spoken French: Experimental evidence. *Journal of French Language Studies*, 24(01), 83-105. 10.1017/S0959269513000380

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Prefixal agreement and impersonal ‘il’ in Spoken French:

## Experimental evidence

Jennifer Culbertson<sup>a, b</sup> and Géraldine Legendre<sup>b</sup>

<sup>a</sup> George Mason University, <sup>b</sup> Johns Hopkins University

### 1. Introduction

Spoken Continental French exhibits a number of properties which were classically argued to characterize null subject languages (Rizzi, 1986b; Jaeggli and Safir, 1989). While it has been shown (e.g. see Gilligan, 1987; Newmeyer, 2005: 45) that this so-called cluster of properties does not characterize all such languages, Spoken French for example appears to allow non-referential (expletive) null subjects and ‘free’ inversion, and lacks *that*-trace effects; (1)a and (1)b show non-referential null subjects in main and embedded clauses, (1)c shows inversion (or right-dislocation) of the subject DP, and (1)d shows the lack of *that*-trace effects.<sup>1</sup>

- (1) a. Ø Faut pas sortir par le four, hein?  
b. J(e) crois que Ø y a Mamie qui arrive.  
c. Elle va dessus (,) la vache.  
d. [<sub>CP</sub> Qui<sub>i</sub> c'est<sub>j</sub> [<sub>IP</sub> t<sub>j</sub> [<sub>CP</sub> t<sub>i</sub> qu(i) [<sub>CP</sub> t<sub>i</sub> est arrivé ?]]]]]

In this paper we focus on constructions like that shown in (1) above, in which the subject clitic *il* (3<sup>rd</sup>, sg. masc.) has been dropped preceding an impersonal verb—here we use this term to refer to any

---

<sup>1</sup> Note that in (1d), the change of complementizer from *que* to ‘agreeing’ *qui* (Rizzi, 1990: 56) can be understood as a repair strategy which does not affect the basic syntax. This and other examples are from the Lyon corpus (see note 2 for details).

verb occurring with a less than fully referential subject. The availability of missing subject clitics<sup>2</sup> in constructions like the impersonal in the Spoken register (where in Standard French, the written or formal register, an overt expletive would be required) has been noted for some time, but has not been investigated in detail (e.g. Lambrecht, 1981: 27-28; Auger, 1993: 179; Coveney, 2002: 30; Fonseca-Greber, 2004). We focus here on providing (i) a clearer picture of what factors affect when a dropped subject clitic in these constructions is accepted by speakers, and (ii) how this phenomenon might be best analyzed within a view of Spoken French subject clitics as *agreement affixes* rather than syntactic subjects (e.g. Roberge, 1990; Auger, 1994; Culbertson and Legendre, 2008; Culbertson, 2010). This view entails that Spoken French is indeed a null subject language, but casts examples like (1a) in a different light—namely as missing an overt agreement morpheme rather than an overt subject. We discuss how Spoken French might fit into a theory of null subject language types cross-linguistically, and show that both usage-based (Bybee and Hopper, 2001) and potentially formally derived (Rizzi, 1986a) factors appear to play a role in explaining when the clitic can be dropped.

It must be noted at the outset that our investigation of this phenomenon, which we refer to as *il*-drop, represents a first pass rather than an exhaustive search of the contexts and factors we believe may affect its availability for speakers of Spoken French. The particular constructions we focus on, however, are motivated by the goals stated above; first, we test constructions that vary in their frequency in the language, second, we test constructions that vary in the *type* of expletive subject licensed—quasi-argumental vs. fully non-argumental—as classified by Rizzi (1986a). Thus we include impersonal verbs as in (1d), but also weather predicates and a number of other distinguishable types (described in detail below).

We proceed as follows. In section 2 we provide a summary of the evidence that subject clitics behave as agreement affixes in Spoken French (in contrast to Standard French). In section 3 we discuss

---

<sup>2</sup> In addition to *il*, the clitic *ça* (sg. demonstrative) is used (and can be dropped) in some constructions licensing a non-referential subject. Here we focus on *il*, but discussion of *ça* is included where relevant.

the relevant constructions in which *il*-drop may be found, and how this phenomenon can be seen in relation to the analysis of subject clitics in Spoken French put forward in section 2. In section 4 we report the results of a controlled acceptability judgment task investigating the factors that influence the availability of *il*-drop. In section 5 we suggest an analysis of *il*-drop which follows from featural matching requirements of subject clitic agreement (Suñer, 1992; Culbertson, 2010), and spell out future directions for research on this topic.

## 2. Spoken French subject clitics as agreement affixes

Traditional analyses of subject clitics in (Standard) French treat them as phonological clitics, attached to the verb in the phonology but functioning as true subject arguments in the syntax (e.g. Kayne, 1975; Rizzi, 1986b; De Cat, 2007). This view is largely based on differences between French and Northern Italian dialects (NIDs), various features of subject clitics in the latter having been argued to follow from their status as agreement affixes rather than true subjects (Rizzi, 1986b; Brandi and Cordin, 1989). In particular, the properties shown in 1–4 below have been presented as evidence for the agreement status of subject clitics in the relevant dialects (see e.g. Rizzi, 1986b, Brandi and Cordin, 1989, De Cat, 2007, and others).

1. No non-affixal material intervening between subject clitic and verb.
2. No (syntactic) movement of the clitic independent of verb host.
3. Obligatory repetition (narrow scope) of subject clitic in conjoined verb phrases.
4. Co-occurrence of subject clitic and DP subject without dislocation/topicalization of the DP.

Note that properties 1–3 pertain to whether a given subject clitic should be analyzed as a phonological clitic or affix, while property 4 pertains to whether that clitic *functions* as an agreement marker. (A clitic with properties 1–3 but not 4 could be treated as a morphologically bound or

incorporated argument, as has been claimed for French object clitics; Miller, 1991). Culbertson (2010) shows that all the above properties in fact hold of Spoken Continental French, just as they do for the NIDs studied by Brandi and Cordin (1989) and others.

Evidence concerning property 1 above typically focuses on the position of negation in Standard French (2)a compared to the NIDs (2)b . In the Standard French example, *ne* intervenes between the subject clitic and the verbal complex, while in Trentino and Fiorentino the negative element does not intervene (abstracting away from some variation in ordering in Fiorentino, Brandi and Cordin, 1989: 120). These data have been used to argue that the structural position of the subject clitic in French must be relatively high compared to the NIDs, i.e. in the same position as other subjects (Rizzi, 1996b; Zanuttini, 1997: 30-35; Poletto, 2000: 18-19), or alternatively that the subject clitic must not attach to the verb pre-lexically (as affixes do, De Cat, 2007: 15-18). However, in Spoken French, the negative marker *ne* is in fact dropped preferentially just when it would intervene between the subject clitic and verb. In the corpus surveyed in Culbertson (2010)<sup>3</sup>, the retention rate in this context is a mere 6.3%, compared to 83.3% when only a subject DP is present.<sup>4</sup>

(2) a. Tu ne le lui as pas dit. (Standard French)

b. Tu le lui as pas dit. (Spoken French)

c. No te ghe l'ha dit. (Trentino)

not you to-him it has said

d. Un tu gliel'ha detto. (Fiorentino)

not you to-him it has said

---

<sup>3</sup> The Lyon corpus (Demuth and Tremblay, 2008) is a large (~50,000 word) corpus of child and child-directed speech. Culbertson (2010) analyzed the child-directed speech in this corpus, a clear source of the colloquial Spoken French register.

<sup>4</sup> The idea that absence of negative *ne* allows the subject clitic to attach morphologically to the verb is discussed in Culbertson (2010: 95), as well as in earlier studies of French negation (Ashby, 1981: 680; Compennolle, 2008: 324; Coveney, 2002: 72-73; Hansen and Malderez, 2004: 15).

Similar conclusions have been drawn based on the existence of subject clitic inversion in question formation in French, where the subject clitic could be analyzed as moving independently of the verb in the syntax (property 2). The examples in (3) illustrate three strategies for question formation, only one of which, (3)a , involves inversion. The latter strategy is not in fact available for subject clitics in the Spoken register (only a residue of 0.9% of inversions were found in the corpus surveyed in Culbertson 2010; see also Coveney, 2002 for additional discussion of question formation in Spoken French).

- (3) a. *Ecoute-t-il?* *(Standard French)*  
       b. *Il écoute?* *(Spoken French)*  
       c. *Est-ce qu'il écoute?* *(Standard, Spoken French)*  
       ‘Is he listening?’

As for property 3 above, Rizzi (1986b) and others point out that in conjoined VPs, the subject clitic in the NIDs must be repeated, as expected of an agreement marker (4)c . By contrast, in Standard French the subject clitic may take wide scope over the conjoined phrases, appearing only once (4)a, suggesting again that they are structurally high and/or must not be attached pre-lexically as an affix. However, in Spoken French, the repetition of the subject clitic is obligatory as shown in (4)b (repeated in 98.4% of cases in the corpus surveyed in Culbertson, 2010) as in the NIDs.

- (4) a. *Elle chante et (elle) danse.* *(Standard French)*  
       b. *Elle chante et \*(elle) danse.* *(Spoken French)*  
       c. *La canta e \*(la) balla* *(Trentino)*  
       ‘She sings and (she) dances.’

Finally, the NIDs show regular and in some cases obligatory *subject doubling*, where the subject clitic appears even when a DP subject is also present. Subject doubling occurs in both Trentino and Fiorentino even when the DP subject is quantificational in nature, as in (5)d , therefore the doubling construction cannot be analyzed as a dislocation (Rizzi 1986b). In Standard French, a similar construction is possible (5)a, however quantificational subjects cannot be doubled by a subject clitic (5)c, suggesting a dislocation analysis. In Spoken French, while still impossible with quantificational subjects, doubling as in (5)b is strikingly frequent—obligatory with all tonic pronoun subjects, and present in 81% of sentences with non-pronominal DP subjects in the corpus surveyed in Culbertson (2010) (see Nadasdi, 1995, Coveney, 2005, among others for quantification of subject doubling rates with particular noun phrase types in other corpora). Further, acoustic analysis of sentences with doubling provides no evidence for the prosodic break that accompanies clear cases of (e.g. object) dislocation, and doubling is possible in broad focus constructions independently analyzed as simple IPs (Culbertson, 2010: 113, contra De Cat, 2004). This places Spoken French squarely within the typology of doubling shown in **Error! Reference source not found..**

- |     |    |                                |                                  |
|-----|----|--------------------------------|----------------------------------|
| (5) | a. | Jean, il a rien dit.           | <i>(Standard French)</i>         |
|     | b. | Jean/lui il a rien dit.        | <i>(Spoken French)</i>           |
|     | c. | Personne (*il) a rien dit.     | <i>(Standard, Spoken French)</i> |
|     | d. | Nessuno *(gl') ha detto nulla. | <i>(Fiorentino)</i>              |
- ‘Jean/he/nobody has said anything.’

[Insert Table 1 here]

### 3. Spoken French as a null subject language and *il*-drop

The evidence summarized above does not support a clear distinction between subject clitics in NIDs like Trentino and Fiorentino and those in Spoken French. Subject clitics have progressed along a grammaticalization cline (from independent pronouns) to the point where they function as agreement markers, attached to the verb in the morphology. A natural consequence of this analysis is the null subject status of these languages—clearly sentences which have a subject clitic agreement alone must now be treated as licensing an empty (null) subject argument.

As mentioned in section 1, Spoken French exhibits a number of other features commonly found in Romance null subject languages, including constructions (6)a which on the surface parallel the non-referential null subjects required for example in Italian. Indeed, dropping expletive *il* in *Standard* French (where subject clitics are arguably still argumental), would be equivalent to dropping an argumental subject. However, if the clitic *il* has become an agreement marker in Spoken French, then what is missing in these sentences is not an expletive subject but (prefixal) agreement. This observation is crucial to understanding how Spoken French might fit into a typology of null subject languages, since under this view both referential and non-referential null subjects are allowed, however certain constructions which take a non-referential null subject do not trigger subject clitic agreement, resulting in *il*-drop. This lack of subject clitic where it would agree with a non-referential null subject has also taken place in the NIDs—in Trentino no subject clitic agreement is present (6)b, in Fiorentino the neuter form appears (6)c.

(6) a. Ø Faut que Mario (il) parte. (Spoken French)

b. Ø Bisogna che el Mario el parta (Trentino)

c. E bisogna che Mario e' parte (Fiorentino)

scl-neuter necessary that Mario scl-masc leave

‘It is necessary that Mario leave.’



How this agreement pattern might be best understood in Spoken French (and the NIDs) will be taken up in section 5. First however, we will discuss variation in the availability of *il*-drop in Spoken French and report the results of an investigatory acceptability judgment task. As noted by a number of researchers, some impersonal predicates allow widespread *il*-drop, while others do so to a lesser extent or not at all (Auger, 1993: 179; Fonseca-Greber, 2004). These contrasts are illustrated in (7) below (examples taken from the Lyons and York CHILDES corpora; Plunkett, 2002), which impressionistically proceed from more to less frequent *il*-drop. (Note that reduction of *il* to *i* is typical in the Spoken register, and possible in all these contexts).

- (7) a. Ø Faut bien le mélanger. *(frequent omission)*  
       ‘It’s necessary that we mix it.’
- b. Ø Y a des livres ici. *(frequent omission)*  
       ‘There are some book there.’
- c. T’as vu Ø en reste encore hein. *(attested omission)*  
       ‘You saw there are more left there huh.’
- d. Et là, i reste plus qu’un croissant. *(more common retention)*  
       ‘And there, there’s more than one croissant left.’
- e. Il s’agit de réparer la boîte. *(common retention)*  
       ‘It’s about fixing the box.’
- f. I pleut un p(e)tit peu. *(retention)*  
       ‘It’s raining a little bit.’

Based on corpus examples from Spoken French like those above, one possibility is that frequency is the driving factor behind the availability of *il*-drop. This is consistent with frequent impersonal constructions such as *(il) faut...* ‘it is necessary...’ and *(il) y a* ‘there is...’ which allow *il*-drop pervasively (see e.g. Fonseca-Greber, 2004) while others do only occasionally or not at all. However, it is also possible that even after controlling for frequency, particular predicates may exhibit differences in the availability of *il*-drop. For example, weather verbs are often analyzed as subcategorizing for a quasi-argumental subject rather than a true expletive, as discussed below (Rizzi, 1986a: 528-529; Travis, 1984: 219; Bennis, 1986: 95-97; Vikner, 1995: Ch. 7). This may explain the apparent unavailability of *il*-drop with weather predicates like *pleuvoir* ‘to rain’ and *faire beau* ‘to be nice (outside)’ (precisely why this prediction would hold is discussed in section 5). Further, in her discussion of the typology of impersonal constructions, Travis (1984) differentiates among a more diverse set of verb classes, suggesting that additional finer-grained categories may also be relevant. These factors are difficult to test based on corpora as (aside from a few frequent ones) impersonal constructions on the whole are quite rare. By contrast, statistical analysis of experimentally gathered native speaker judgments makes it possible to vary factors of interest in a controlled manner, and directly contrast particular impersonal sentences with and without *il*.

Below we investigate the effect of frequency, non-argumental vs. quasi-argumental status, and also verb class (explained in detail below) on native speakers’ ratings of impersonal sentences with and without *il*.

#### 4. Acceptability ratings of *il*-drop by native speakers

##### 4.1. Participants

Participants were 24 native speakers of Continental French recruited through Amazon Mechanical Turk, an online crowdsourcing marketplace which pairs companies or individuals posting tasks with individuals willing to complete them (see Sprouse, 2011 for validation of this service as a methodology

for syntactic experiments). Participants ranged in age from 18-79, and were from various regions in France (including Paris, Lyon, Bordeaux, Bretagne). To ensure that participants were indeed native speakers of French, IP addresses were constrained to within France, and a test of fluency was given (described below).

#### 4.2. Stimuli

Impersonal predicates in French are best separated into four classes combining semantic properties and syntactic considerations: weather predicates, modals, non-modals and ‘existential’ predicates. As described below the variety of syntactic properties associated with impersonal predicates argues against a purely syntactic classification (cf. Travis, 1984). The first class contains weather verbs as in *il neige* ‘it is snowing’ and *il pleut* ‘it is raining’, and also adjectival predicates like *il fait beau* ‘it is nice’ and *il fait du vent* ‘it is windy’. Expletive *il* in weather verb constructions is held to be a quasi-expletive, since it can serve as a controller of (non-arbitrary) PRO in a non-finite complement clause, as shown in (8a,b) (Chomsky, 1981: 323-325; Rizzi, 1986a: 529). By contrast, expletive *il* in all other target constructions involving a non-finite complement clause may not control PRO, and examples such as (8d) necessarily involve *arbitrary* reference rather than co-reference.

- (8) a. It<sub>i</sub> rained before PRO<sub>i</sub> snowing.  
b. Il<sub>i</sub> a plu avant de PRO<sub>i</sub> neiger.  
c. It is enough PRO<sub>arb</sub> to talk.  
d. Il suffit de PRO<sub>arb</sub> parler.

The second class contains epistemic and non-epistemic modals as in *il faut...* ‘it is necessary...’, *il vaut mieux...* ‘it is better...’, *il est possible...* ‘it is possible...’, which take a sentence complement introduced by *que* or in some cases an infinitival verb (e.g. *il vaut mieux qu’on parte* ‘it is better that we leave’; *il vaut mieux partir*, ‘it is better to leave’). Epistemic modals may undergo extraposition

optionally ([<sub>TP</sub> *il est possible/probable* [<sub>CP</sub> *que...*]] vs. [<sub>TP</sub> [<sub>CP</sub> *que...*] *est possible/probable*] ‘it is possible/likely that...’ vs. ‘that...is possible/likely’) or obligatorily ([<sub>TP</sub> *il se peut* [<sub>CP</sub> *que...*]] vs. \*[<sub>TP</sub> [<sub>CP</sub> *que...*] *se peut*], ‘it may be that...’ vs. ‘\*that...may be.’). The predicates in the third class typically take sentential complements, but are not modal, e.g. *il semble/paraît que...* ‘it seems that...’, *il s’agit de...* ‘it’s about...’, *il m’arrive de...* ‘it happens to me...’. Some are passive (*il est prévu que...* ‘it is expected that...’, *il est convenu que...* ‘it is agreed that...’). The fourth class contains intransitive predicates which may appear in an impersonal existential (or presentational) construction subject to a definiteness restriction (see e.g. Labelle, 1992: 381-382; Cummins, 1996: Ch. 2.5): *il y a DP<sub>-def</sub>* ‘there is DP<sub>-def</sub>’, *il reste DP<sub>-def</sub>* lit. ‘there remain DP<sub>-def</sub>’, *il sort DP<sub>-def</sub>* lit. ‘there leave DP<sub>-def</sub>’, etc. Following Lambrecht (1994), Legendre and Sorace (2010) divide these predicates into presentational and event-reporting on the basis of their discourse properties: presentational predicates serve to introduce a new, not yet activated, entity in the world of discourse while event-reporting predicates merely introduce a new event. Both varieties are grammatical with all lexico-aspectual classes in French, including those traditionally associated with unergativity (9).

- (9) a. Il gisait un homme sur le trottoir. (Presentational)  
       ‘There lay a man on the sidewalk.’  
       b. Il a échappé une bourde au président. (Event-reporting)  
       ‘There slipped a gaff from the president.’

Test sentences (42 total) were constructed using a sample set of predicates from each class (listed in Table 2 below). The frequency of each predicate was estimated based on use in impersonal constructions in the PFC corpus (Phonologie du français contemporain; Durand, Laks and Lyche, 2002), an adult-directed corpus of Spoken French. Predicates were separated into three frequency

bins—high, mid, low.<sup>5</sup> Frequency information for each verb is shown in **Error! Reference source not found.** A number of these verbs can also appear with *ça* in the same constructions, for example *ça suffit*, and *c'est clair* are possible, but not *\*ça faut* or *\*ça y a*.<sup>6</sup> Thus *ça* functions as an alternative marker in impersonal constructions in Spoken French, perhaps unsurprisingly as it also alternates with *il* in referential contexts particularly when the subject is generic (*ça* can be treated as an agreement marker with a potentially distinct featural makeup; Auger, 1994: Ch. 4; Culbertson, 2010: 123). For this reason then, the absence of *il* in at least some of our sentences could also be interpreted as absence of *ça*. We return to this issue below, however note that what we are interested in here is whether an agreement marking clitic can be dropped.

[Insert Table 2 here]

Using the stimuli described above, four lists were created, comprised of 22 test sentences and 25 fillers each. In half of the test sentences *il* was dropped, and in the other *il* was retained. Approximately half of the filler sentences were expected to be rated as ungrammatical and the other half as grammatical. Each participant judged sentences from only one of the four lists.

Sentences were recorded by two young native French speakers (one female, one male) who knew one another and were instructed to use the informal register (this was encouraged in the orthography of the script the speakers read from as well, e.g. *ne* was omitted, *t'* (2<sup>nd</sup>, sg. acc.) was used

---

<sup>5</sup> Only a subset of the PFC corpus was used; search was restricted to speakers in France, and only utterances classified as ‘discussion libre’ were queried in order to get a sample representative of Spoken French. (The PFC corpus was used here instead of a child-directed speech corpus since many of the verbs tested are not used in speech to children). Low frequency corresponded to 0-19 attested uses in impersonal constructions, mid corresponded to 20-79, high corresponded to greater than 80 (the highest frequency construction was *il y a* ‘there is’ with 4969 instances).

<sup>6</sup> Exactly which can appear with *ça* may differ across varieties of French; for example an anonymous reviewer and speaker of Québec French reports that *ça* is not possible with weather predicates in her grammar. However for the second author, a speaker of European French, these are perfectly acceptable.

instead of *te* before a consonant, etc.). For each test sentence, versions with and without *il* were recorded. The two speakers were allowed to practice beforehand to ensure fluency for each sentence. Context was provided in the form of a set-up question, as in the example in (10) below. Note that each rated sentence (both test and filler) was preceded by either *oui*, *non* or another conjunction so that the verb (accompanied by *i(l)* or not) would not be sentence-initial.

- (10) Q: Tu veux t'promener ?                      (*set-up*)  
      'You wanna go for a walk?'  
      A: Oui, i(l) fait beau ce soir.              (*rated sentence*)  
      'Yes, it's nice out tonight.'

#### 4.3. Procedure

Participants were recruited through Amazon Mechanical Turk, and were first required to answer a number of demographic questions concerning their age, sex, birthplace, current residence, and other language experience. They were asked to verify that (i) they had lived in France from birth until at least age 15, and (ii) their parents had spoken French to them during that time. They also completed a test of their fluency which required them to listen to a clip of a French song and briefly describe it.

Participants were then instructed that they would hear auditorily presented sentence pairs, and their task was to rate the *second* sentence (the answer to the preceding question). They were explicitly told that experimenters were interested in their judgments of the informal register.<sup>7</sup> Sentences were rated on a scale from 1 (*totalement inacceptable* 'totally unacceptable') to 7 (*tout à fait acceptable* 'perfectly acceptable'), intermediate points were not labelled (following Schütze, 1996: 189, note 12). Participants pressed play to hear each question-answer pair, and then entered their rating.

---

<sup>7</sup> The text of the instructions can be found in the appendix.

#### 4.4. Results

As a sanity check, we first verified that participants rated filler sentences expected to be grammatical as higher than those expected to be ungrammatical. Data for this and all other analyses presented here were subjected to mixed effects logistic regression, which allows for the combination of fixed and random effects in a single model.<sup>8</sup> Here we treat participant as a random effect. As expected, participants reliably rated grammatical fillers higher than ungrammatical fillers (mean rating 5.45 ( $sd = 1.87$ ) vs. 3.92 ( $sd = 2.36$ ),  $\beta = 1.53 \pm 0.13$ ,  $p < 0.001$ ).

Setting aside filler items, we predicted that increased frequency would lead to higher ratings for test sentences in which *il* was dropped. To test this, we ran a 2 x 3 model with sentence type (*il*, no *il*) and frequency (high, mid, low) as predictors. The results are summarized in **Error! Reference source not found.** below.

[insert Table 3 here]

First, there was a marginally significant effect of sentence type, with participants rating sentences with *il* retained somewhat higher than those with *il* dropped. This is as expected if French speakers have access to both the Standard and Spoken register, and retaining *il* is clearly (in general) an acceptable possibility in both (Massot, 2008; Culbertson, 2010: 97; Palais, 2013). Participants were also found to generally rate sentences containing low frequency verbs lower than those with high frequency verbs.<sup>9</sup> Most interestingly, a significant interaction between sentence type and frequency was found—participants rated sentences with *il* dropped significantly higher when those sentences contained a high frequency verb compared to a mid and low frequency verb. To investigate this effect further, sentences with *il* dropped were analyzed alone. When *il* was dropped, participants were found

---

<sup>8</sup> All p-values reported here were estimated using MCMC sampling from the posterior distribution (implemented in `pvals.fnc()` in the `languageR` package for R statistical software; Baayen 2008).

<sup>9</sup> Frequency is often found to affect acceptability judgments (Schütze, 1996: 164), however in this case the effect might also be driven by the fact that lower frequency verbs are also not typically used in the Spoken register.

to rate sentences with high frequency verbs higher than those with mid frequency ( $\beta = -1.50 \pm 0.26, p < 0.001$ ), and higher than those with low frequency ( $\beta = -1.58 \pm 0.20, p < 0.001$ ). No difference was found between ratings for sentences with mid and low frequency verbs ( $\beta = 0.008 \pm 0.10, p = 0.67$ ). In summary, when the verb in the sentence was high frequency, participants in this experiment accepted sentence with *il*-drop almost as readily as they accepted their *il*-retained counterparts (mean 5.58 compared to 6.01). By contrast, for sentences with only mid or low frequency verbs, ratings of *il*-drop were relatively low. This result confirms our prediction that frequency of the verb indeed has a strong effect on the availability of *il*-drop in Spoken French. Example sentences (with their average ratings out of 7) are given in (11).

- (11) a. Ben, Ø y a un embouteillage. (High, avg. 6.6)  
       ‘Well, there was a traffic jam.’
- b. Non, Ø me reste encore soixante-deux pages à lire. (Mid, avg. 4.8)  
       ‘No, I still have sixty-two pages to read.’ (lit. ‘there remain to me...’)
- c. Oh, il m'arrive de faire des erreurs parfois ! (Low, avg. 4.4)  
       ‘Oh, it happens that I make mistakes sometimes!’

Verb type—as defined by the verb classes described in section 4.2—was also investigated in more detail for test sentences with *il* dropped. **Error! Reference source not found.** illustrates the average rating given for sentences with verbs in each class. First, to confirm that verb class in fact contributed significantly to explaining the data, we added this as a predictor to our previous model which contained verb frequency as a predictor. A Likelihood Ratio test (Lehmann, 1986) confirmed that the added complexity of the model with verb class as a predictor was warranted by the improved fit to the data ( $\chi^2(3) = 11.81, p = 0.008$ ).



[insert Figure 1 here]

The results of the regression model are summarized in **Error! Reference source not found.** below. This model used Helmert contrast coding, which allows us to evaluate several key contrasts in a single model by testing each level of the verb class factor against the mean of all previous levels. (The frequency factor is coded such that comparisons are made to the mean rather than a single level, since we are focusing here on the effect of verb type.) With *il* dropped, participants rated sentences with non-modal verbs (taking sentential complements), as marginally higher than those with weather verbs. Sentences with weather or non-modal verbs were not rated differently from ‘existential’ verbs. However comparing the latter verb types with the class of modal impersonals reveals that sentences with modal verbs were rated significantly higher. Representative sentences from each class, along with their average ratings (out of 7), are given in (12).

[insert Table 4 here]

- (12) a. Oui, mais Ø neige tout le temps. *(Weather, avg. 3.5)*  
‘Yes, but (it) snows all the time.’
- b. Oh, Ø m'arrive de faire des erreurs parfois ! *(Non-modal, avg. 4.4)*  
‘Oh, (it) happens that I make mistakes sometimes!’
- c. Non, Ø reste deux heures. *(Existential, avg. 5.5)*  
‘No, (there) are two hours left.’
- d. Ouais, Ø vaut mieux y aller en plein jour avec plusieurs potes. *(Modal, avg. 6.6)*  
‘Yeah, (it) is better to go there in the daytime with a bunch of friends.’

Interestingly, additional post-hoc comparisons (Bonferroni corrected) revealed that among sentences with weather predicates, *il* drop was rated significantly higher with adjectival predicates like ‘fait beau’ compared to verbs like ‘pleut’ ( $\beta = -1.75 \pm 0.58, p = 0.002$ ). The highest and lowest rated examples in the weather class shown in (13) are representative of this difference (note example (13)a is in the range of acceptability of *il*-drop with predicates like *s’agit* ‘is about’ and *semble* ‘seems’).

- (13) a. Non,  $\emptyset$  faisait un peu froid. (Adjectival, avg. 5.4, cf. with *il* 6.2)  
           ‘No, (it) was a bit cold.’  
       b. Non,  $\emptyset$  pleut. (Verbal, avg. 2.0, cf. with *il* 6.5)  
           ‘No, (it) is raining.’

Finally, as suggested by an anonymous reviewer, we investigated two additional effects. First, the effect of phonological context on *il*-drop ratings. In particular, we looked to see whether *il*-drop was more likely to be accepted when *il* would be followed by a vowel-initial auxiliary verb or verbal complex (e.g. a auxiliary verb like *être* ‘to be’, or a clitic+verb complex as in *y a* ‘(there) is’). This factor was not found to be significant ( $\beta = -0.20 \pm 0.20, p = 0.32$ ), however as it was not systematically manipulated in our stimuli, additional investigation may be warranted. Second, we conducted several analyses aimed at assessing whether particular verbs behaved differently than others within a given verb type (weather, modal, non-modal, existential). Among the weather verbs, sentences using *faire froid* ‘to be cold’ were found to be more acceptable without *il* than other weather predicates, while sentences using *pleuvoir* ‘to rain’ were found to be less acceptable. Among the modal verbs, sentences with *vaux mieux* ‘to be better’, and *être probable* ‘to be possible’ were found to be more acceptable without *il*, while *être clair* ‘to be clear’ was less acceptable. Among the non-modal

verbs, sentences with *suffir* ‘to be enough’ were found to be more acceptable without *il*, while sentences with *prevoir* and *convenir*, e.g. *est prévu que...* ‘is predicted that...’, *est convenu que...* ‘is agreed that...’ were less acceptable. Finally among existential predicates, sentences with *venir*, e.g. *il vient DP-def* ‘there comes DP-def’ were found to be more acceptable without *il*, while those with *dormir*, e.g. *il dort DP-def* ‘there sleeps DP-def’ was found to be less acceptable. (In all cases,  $p < 0.05$ , and models are fit on residual ratings after frequency is taken into account). This suggests that there may be some amount of construction-specific variation in addition to variation between classes of verbs, however further investigation would be needed to confirm these differences.<sup>10</sup>

## 5. Discussion and conclusions

### 5.1. Summary

Based on impressionistic analysis of corpus data, we hypothesized that frequency and quasi-argument vs. true expletive status of the impersonal subject are factors that affect the availability of *il*-drop in Spoken French. We predicted that higher frequency verbs would result in higher ratings of sentences with *il*-drop, and that verbs taking a quasi-argument (e.g. weather predicates) would result in lower ratings of *il*-drop. These predictions were borne out by the results of the acceptability judgment task reported here. In particular, sentences with *il*-drop which used verbs in the highest frequency bin (e.g. *(il) faut (que)...* ‘(it) is necessary (that)...’, *(il) y a...* ‘(there) is...’, *(il) paraît que...* ‘(it) seems that...’) were rated higher compared to sentences using mid or low frequency verbs, in line with frequency-based models of change (e.g. Bybee and Hopper, 2001). Crucially however, frequency alone did not fully explain participants’ ratings. First, *il*-drop in sentences with weather predicates was generally rated lower than for other verb types. Highest rated with *il*-drop were sentences using verbs we have classified here as modals (which often take a sentential complement), e.g. *faut (que)*, *est*

---

<sup>10</sup> The anonymous reviewer also suggests that there may be some relationship between the ability of a construction to appear with *ça* in addition to *il* and the availability of dropping either one. Based on our data this does not appear to be the case—verbs which cannot take *ça* include both those rated highly without *il*, as in *(il) faut...*, and those given low ratings, as in *(il) sort...*

*possible/clair/probable (que)*, *vaut mieux (que)*. The other verb classes tested—non-modals (with sentential complements) and presentational/event-reporting verbs—had ratings which fell somewhere in the middle. Again, these asymmetries held even after controlling for the effects of frequency. The picture painted by these results is a complex one, and likely represents a snap-shot of a change in progress; there is without a doubt variation in *il*-drop even for those predicates with which it was judged most acceptable here, and surprisingly, even predicates typically thought to categorically disallow *il*-drop were not necessarily rejected by speakers (e.g. adjectival weather predicates).

## 5.2. Discussion

We have established, using controlled native speaker judgments, that Spoken Continental French does indeed systematically allow the subject clitic agreement *il* to be dropped in certain impersonal constructions. We now return to the issue of what this means for the grammar of the language, and *why* it might be available in some contexts and not others. First, as mentioned in section 3, this loss of subject clitic agreement in the context of impersonal constructions has also taken place in the NIDs (examples repeated in (16)). This process therefore seems likely to be a reflection of the kind of grammaticalization which has occurred in these languages. While these constructions on the surface are parallel to the non-referential null subjects found in Italian and other null subject languages, there is an important distinction; in Italian for example, just as in referential null subject contexts, there is no overt subject *argument* in impersonal constructions, however in both cases suffixal verbal *agreement* marking is present. In Spoken French and the NIDs, an overt subject argument can be dropped in referential and non-referential contexts, *however*, the prefixal agreement marking can also be dropped in (some) non-referential impersonal constructions. Thus, because the subject clitics in these languages are not true argumental subjects, dropping *il* (or the equivalent clitic in the NIDs) is parallel to dropping an *agreement marker*, not a subject. When the *il* is retained in impersonal constructions, there is still a non-referential null subject. What needs to be understood for these language then, is why

prefixal agreement is not triggered (regularly) in impersonals. We focus first on the difference between the availability of *il* drop in weather predicates compared to the other construction types investigated.

In the literature on null subject languages, the following typology of null pronominals has been suggested (Rizzi, 1986a): *pro*<sub>+ref</sub> (fully referential/argumental null pronoun), *pro*<sub>expl+A</sub> (quasi-argumental null pronoun), *pro*<sub>expl-A</sub> (non-argumental null pronoun). These distinctions are motivated by asymmetries in the availability of null subjects of these types cross-linguistically, and in the *form* of the expletive used (Travis, 1984). Setting aside for the moment the apparent variability in the Spoken French system, we can characterize the language within this typology as licensing all three null pronominals, but exhibiting (obligatory) clitic agreement with only *pro*<sub>+ref</sub> and *pro*<sub>expl+A</sub>. The suggestion we would like to put forward here is that the featural content of these three null pronominals differs, resulting in the difference in agreement triggering. In particular, subject clitic agreement is not triggered by true expletive *pro*<sub>expl-A</sub> precisely because this argument lacks any featural content—it cannot provide an identifiable set of subject phi-features, required for subject clitic agreement. This is similar to what was implied by Lambrecht (1981: 28), who characterized Spoken French subject clitics as fulfilling an agreement function which is crucially not available in impersonal constructions. By contrast, *pro*<sub>+ref</sub> clearly has a full set of phi-features, and thus is expected to trigger clitic agreement. The quasi-argumental *pro*<sub>expl+A</sub>, we suggest, while not having a fully articulated set of features nevertheless has *some* featural content, in fact enough to trigger agreement. Evidence for this view comes from several sources which we survey briefly here (for the most part based on overt quasi-arguments and expletives, but by assumption the properties which hold of these also hold of the null *pro* alternatives).

First, as mentioned above, quasi-argumental expletives differ from true expletives in serving as controller of (non-arbitrary) PRO (8), have been argued to carry a bleached (rather than no) theta-role (Chomsky, 1981: 323-325; Rizzi, 1986a: 528-529), and have some (rather than no) referential value (Bolinger, 1977: 77-87; Rizzi, 1986a: 528-529). Along similar lines, Svenonius (2002: 6) points out

that because weather predicates can assign case to an internal object (‘it’s raining mackerel/cats and dogs/men, *il pleut des cordes/grenouilles*), by Burzio’s Generalization, they must actually assign an external theta-role to ‘it’ in English (although see Crousaz and Benincà, 2003 on how this may vary cross-linguistically). Acquisition evidence also supports a referentiality distinction between *pro<sub>expl+A</sub>* and *pro<sub>expl-A</sub>*. In particular, Kirby and Becker (2007) show that English children acquire referential and quasi-argumental ‘it’ earlier than true expletive ‘it’, and interpret this finding as support for the view that children acquire more referential arguments earlier. Cross-linguistically, special ‘dummy’ subjects e.g. ‘world’, ‘nature’, ‘surroundings’, or cognate arguments ‘rain (N)’, ‘snow (N)’ can be used in weather constructions but not in true expletive constructions (Eriksen, Kittilä and Kolehmainen, 2010). In French (like English), most weather predicates exhibit the alternations in (14) not found in other impersonal constructions (although for more details see the thorough discussion in Ruwet, 1991: Ch. 3). It is possible then that the referential accessibility of quasi-argumental expletives comes in part by analogy to such constructions, or to something like incorporation of the subject/theme *la pluie* into the verb (as suggested by Jackendoff, 1983: 185).

- (14) a. Il pleut aujourd’hui.  
           ‘It’s raining today.’
- b. Il tombe de la pluie.  
           Lit. ‘It is falling (of the) rain.’
- c. La pluie tombe aujourd’hui.  
           ‘The rain is falling today.’
- d. Le temps est pluvieux aujourd’hui.  
           ‘The weather is rainy today.’

A more precise characterization of the feature difference between *pro<sub>expl+A</sub>* and *pro<sub>expl+A</sub>* was suggested by Rizzi (1986a: 543), who argued that quasi-arguments are in fact specified for *number features* while true expletives have neither number nor person features (see also Ouhalla, 1993: 490).

That the presence of subject clitic agreement would be sensitive to the features of the subject—either null or overt—is perhaps not surprising given other facts about Spoken French and the NIDs. Suñer’s (1992) Matching Hypothesis aims to explain the pattern of subject clitic agreement in Trentino and Fiorentino as due to feature-matching requirements which must hold between the subject argument and an agreeing clitic. In particular, she argues that subject clitics in the NIDs are triggered only when the subject argument is given or identifiable (and null or neuter otherwise). This explains, for example, the pattern of missing/default agreement with *wh*-subjects (as in (15)b) also discussed by Brandi and Cordin (1989). Similarly, because the *pro<sub>expl-A</sub>* is not identifiable, in Trentino no subject clitic agreement should appear, and this is confirmed by (16)b. In Fiorentino, there is in fact a “default” neuter clitic, and it is the one used with impersonal verbs (16)c.

Following Suñer (1992), Culbertson (2010) analyzed Spoken French subject clitics as agreement markers triggered only when the subject argument is definite and accessible (see also Nadasdi, 1995, Coveney, 2005). Thus no clitic agreement is present with indefinite DPs (e.g. *personne* ‘nobody’) or *wh*-subjects, as in (15)a. If our discussion of the features of quasi-argumental expletives is on the right track, the marker *il* would be triggered in such cases because the featural set of *pro<sub>expl+A</sub>* does not clash with the matching requirements of subject clitic agreement. On the other hand no subject clitic agreement is predicted to appear with *pro<sub>expl-A</sub>* either because—as Suñer (1992) argued—such subjects are essentially –accessible, or alternatively because no features are present to be matched.<sup>11</sup> As mentioned above, the clitic *ça*, is used as an alternative to *il* in some impersonal constructions,

---

<sup>11</sup> This would of course distinguish suffixal agreement morphology from prefixal subject clitic agreement in Spoken French—suffixal agreement is default 3<sup>rd</sup>-sg-masc in these contexts, while no default (matching) form of the clitic exists and thus it is simply absent. In fact, in Old French, suffixal agreement with expletives also exhibited variation—the agreement could target a post-verbal (nominative) DP or could appear as default 3<sup>rd</sup> singular (Arteaga, 1994).

including with weather predicates (e.g., *ça pleut*). This clitic is argued by Auger (1994: Ch. 4) to be an agreement marker with features distinct from *il* based on its use with generic (referential) subject DPs. In at least some of the constructions tested here, it may be that *ça* rather than *il* is dropped, however in either case clitic agreement is missing. It remains to be seen whether the pattern of omission of *ça* differs from that of *il*.

- (15) a. Qui  $\emptyset$  est venu avec toi? (Spoken French)  
       b. Chi  $\emptyset$  è vegnù con ti? (Trentino)  
       c. Chi **gl'** è venuto con te? (Fiorentino)  
       ‘Who has come with you?’

- (16) a.  $\emptyset$  Faut que Mario (il) parte. (Spoken French)  
       b.  $\emptyset$  Bisogna che el Mario el parta (Trentino)  
       c. **E** bisogna che Mario e’parte (Fiorentino)  
       ‘It is necessary that Mario leave.’

The remaining factors we have uncovered here—that modal impersonal predicates are rated as the most acceptable with *il*-drop, and that there is a clear relationship between frequency and *il*-drop—may be related to more general principles of language change. For example, one of the most frequent impersonal constructions, *faut (que)*, is a modal predicate, and the availability of *il*-drop more generally in that class may be the result of analogical processes (Hopper and Traugott, 2003). Further, the general phonological reduction of *il* to *i* in Spoken French seems like to be similarly related to frequency, and thus may serve to accelerate *il*-drop in just those cases.



While we have couched our explanation here in terms of abstract features, it is possible to formulate an alternative view of this process of grammaticalization which is similar in spirit, but more in line with usage-based approaches (e.g. Bybee and Hopper, 2001; Hopper, 1987). Under this view, subject clitic agreement can in some sense be *expected* to drop in constructions like the ones investigated here; impersonals, weather predicates, and the like present no ambiguity in terms of the semantic identity of the subject, and thus agreement is not functionally needed there. A usage-based approach would of course also lead to the expectation of frequency as a driving factor, and the most frequent constructions—among them those we have classified as modal, non-modal, and existential—indeed are most permissive of *il*-drop. If these constructions can be distinguished from weather verbs semantically, then *il*-drop might extend analogically to the former faster than the latter. Of course, it remains to be worked out precisely how the classes might in fact be distinguished (without recourse to notions like the type of null subject licensed).

Several additional questions remain of course, most obviously the variation in availability of *il*-drop *within* a given impersonal construction in Spoken French. We leave this issue for the most part to future work, however a few comments are in order. The appearance of subject clitic agreement in fact exhibits variation in other contexts in Spoken French as well. In particular, even when the features of an overt argumental subject match those of the subject clitic, agreement is triggered apparently optionally in some cases (e.g. in the case of non-pronominal definite DP subjects, see (16), and (5)b). Thus any analysis of subject clitic agreement in Spoken French will necessarily be faced with this issue. It seems likely that this variation is due to the fact that the clitic system is still in the process of change in French, and most adult speakers by assumption have access to the Standard French grammar—a non-null subject grammar in which subject clitics are true argumental subjects.<sup>12</sup> This is

---

<sup>12</sup> Alternatively, subject clitic agreement is triggered probabilistically within the Spoken French grammar itself (e.g. by stochastic constraint ranking in an Optimality Theoretic framework).

supported by the fact that frequency is clearly playing a role in Spoken French *il*-drop (see Krug, 1998; Diessel, 2007, among others for discussion of the role of frequency in grammaticalization).

### 5.3. Conclusion

Spoken French subject clitics have been argued to function as agreement affixes (Roberge, 1990; Auger, 1994; Culbertson, 2010; among others) similar to the agreement subject clitics of several Northern Italian dialects. These languages have thus become null subject licensing, and have two sources of agreement morphology—an older suffixal system (phonologically retained for verbs of certain conjugation classes only) , and a newer prefixal system. We have focused here on constructions licensing *non-referential* null subjects in Spoken French—impersonal constructions—in which the agreement subject clitic *il* is optionally absent. Using controlled native speaker judgments, we have confirmed that the availability of *il*-drop is affected by the frequency of the predicate, the type of null subject licensed (quasi-argument vs. true expletive), and relatedly, the class of the predicate (modal verbs were given the highest ratings). These findings are a preliminary but significant step forward given the difficulty of investigating these features in-depth using corpus evidence alone. We have argued that the (optional) lack of subject clitic agreement in these constructions, also a feature of the Northern Italian dialects with agreement clitics, is the result of feature matching requirements of these agreement systems (Suñer, 1992; Culbertson, 2010). For Spoken French, we have suggested that a featural (referential) distinction between the quasi-argumental *pro<sub>expl+A</sub>* of weather predicates and the true expletive *pro<sub>expl-A</sub>* of other impersonal constructions leads to (near-)obligatory clitic agreement in the former but not the latter.

### References

- Arteaga, D. (1994). Impersonal constructions in Old French. In M. Mazzola (ed), *Issues and Theory in Romance Linguistics. Selected Papers from the Linguistic Symposium on Romance Languages XIII*. Philadelphia: Benjamins, pp. 141–156.

- Ashby, W. J. (1981). The loss of the negative particle *ne* in French: A syntactic change in progress. *Language*, 57.3: 674–687.
- Auger, J. (1993). More evidence for verbal agreement-marking in colloquial French. In W. Ashby, M. Mithun, and G. Perissinotto (eds), *Linguistic Perspectives on the Romance Languages: selected papers from the XXI Linguistic Symposium on Romance Languages*. Philadelphia: Benjamins, pp. 177–198.
- Auger, J. (1994). *Pronominal Clitics in Québec Colloquial French : A Morphological Analysis*. Ph.D. dissertation, University of Pennsylvania.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Bennis, H. (1986). *Gaps and Dummies*. Dordrecht: Foris.
- Bolinger, D. (1977). *Meaning and Form*. London: Longman.
- Brandi, L., and Cordin, P. (1989). Two Italian dialects and the null subject parameter. In O. Jaeggli, and K. Safir (eds), *The null subject parameter*. Dordrecht: Kluwer Academic Publishers, pp. 111–142.
- Bybee, J., and Hopper, P. (2001). *Frequency and the emergence of language structure*. Amsterdam: Benjamins.
- Chomsky, N. (1981). Markedness and core grammar. In A. Belletti, L. Brandi, and L. Rizzi (eds), *Theory of markedness in core grammar*. Pisa: Scuola Normale Superiore di Pisa, pp. 123–146.
- Compernelle, R. A. van (2008). Morphosyntactic and phonological constraints on negative particle variation in French-language chat discourse. *Language Variation and Change*, 20.2: 317–339.
- Coveney, A. (2002). *Variability in Spoken French: a sociolinguistic study of interrogation and negation*. Exeter: Elm Bank.
- Coveney, A. (2005). Subject doubling in Spoken French: A sociolinguistic approach. *The French Review*, 79.1:96–111.

- Crousaz, I. D., and Shlonsky, U. (2003). The distribution of a subject clitic pronoun in a Franco-Provençal dialect and the licensing of *pro*. *Linguistic Inquiry*, 34: 413–442.
- Culbertson, J. (2010). Convergent evidence for categorial change in french: From subject clitic to agreement marker. *Language*, 86.1:85–132.
- Culbertson, J. and Legendre, G. (2008). Qu'en est-il des clitiques sujet en français oral contemporain? *Proceedings of the 1er Congrès mondial de linguistique française*. Paris, France.
- Cummins, S. 1996. *Meaning and Mapping*. Ph.D. dissertation, University of Toronto.
- De Cat, C. (2007). *French Dislocation. Interpretation, Syntax, Acquisition*. Oxford: Oxford University Press.
- De Cat, C. (2004). On the impact of French subject clitics on the information structure of the sentence. In R. Bok-Bennema, B. Hollebrandse, B. Kampers-Manhe, P. Sleeman (eds) *Romance Languages and Linguistic Theory 2002, Selected papers from "Going Romance", Groningen, 28-30 November 2002*. Philadelphia: Benjamins, pp. 33–46.
- Demuth, K., and Tremblay, A. (2008). Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language*, 35: 99–127.
- Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology*, 25: 108– 27.
- Durand, J., Laks, B., and Lyche, C. (2002). La phonologie du français contemporain: usages, variétés et structure. In C. Pusch, and W. Raible (eds), *Romance Corpus Linguistics - Corpora and Spoken Language*. Tübingen: Narr, pp. 93–106.
- Eriksen, P., Kittila, S., and Kolehmainen, L. (2010). The linguistics of weather: Cross-linguistic patterns of meteorological expressions. *Studies in Language*, 34: 565–601.
- Gilligan, G. M. (1987). *A cross-linguistic approach to the pro-drop parameter*. Ph.D. dissertation, University of Southern California.

- Hansen, A. B., and Malderez, I. (2004). Le *ne* de négation en région parisienne: Une étude en temps réel. *Langage and Société*, 107: 5–30.
- Hopper, P. (1987). Emergent grammar [Electronic version]. *Berkeley Linguistics Society*, 13:139-157.  
Available from: <http://home.eserver.org/hopper/emergence.html>
- Hopper, P. and Traugott E. C. (2003). *Grammaticalization*. Cambridge: Cambridge University Press.
- Jackendoff, R. (1983). *Semantics and Cognition*. Cambridge, MA: MIT Press.
- Jaeggli, O., and Safir, K. (1989). *The Null Subject Parameter*. New York: Springer.
- Kayne, R. (1975). *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.
- Kirby, S., and Becker, M. (2007). Which it is it? the acquisition of referential and expletive *it*. *Journal of Child Language*, 34: 571–599.
- Krug, M. (1998). String frequency: A cognitive motivating factor in coalescence, language processing and linguistic change. *Journal of English Linguistics*, 26: 286– 320.
- Labelle, M. 1990. Change of State and Valency. *Journal of Linguistics* 28: 375-414.
- Lambrecht, K. (1981). *Topic, Antitopic and Verb Agreement in Non-Standard French*. Philadelphia: Benjamins.
- Lambrecht, K. (1994). *Information structure and sentence form*. Cambridge: Cambridge University Press.
- Legendre, G., and A (2010). Auxiliaries and intransitivity in French and in Romance. In D. Godard (ed), *Fundamental Issues in the Romance Languages*. Stanford, CA: CSLI Publications, pp. 171–220.
- Miller, P. (1991). *Clitics and Constituents in Phrase Structure Grammar*. New York: Garland.
- Nadasdi, T. (1994). Subject np doubling, matching, and Minority French. *Language Variation and Change*, 7.1:1–14.
- Newmeyer, F. (2005). *Possible and Probable Languages*. Oxford: Oxford University Press.

- Ouhalla, J. (1993). Subject-extraction, negation and the anti-agreement effect. *Natural Language and Linguistic Theory*, 11: 477–518.
- Palais, K. (2013). The case for diglossia: Describing the emergence of two grammars in the early acquisition of Metropolitan French. *Journal of French Language Studies*, 23: 17–35.
- Plunkett, B. (2000). What’s ‘what’ in french questions. *Journal of linguistics*, 36: 511–530.
- Poletto, C. (2000). *The Higher Functional Field. Evidence from Northern Italian Dialects*. Oxford: Oxford University Press.
- Rizzi, L. (1986a). Null objects in Italian and the theory of pro. *Linguistic Inquiry*, 17: 501–557.
- Rizzi, L. (1986b). On the status of subject clitics in Romance. In O. Jaeggli, and C. Silva- Corvalan (eds), *Studies in Romance Linguistics*. Dordrecht: Foris, pp. 391–419.
- Rizzi, L. (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.
- Roberge, Y. (1990). *The Syntactic Recoverability of Null Arguments*. Montreal: McGill-Queen’s University Press.
- Ruwet, N. (1991). *Syntax and the Human Experience*. Chicago: University of Chicago Press.
- Schütze, C. (1996). *The Empirical Base of Linguistics*. Chicago: University of Chicago Press.
- Sprouse, J. (2011). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods*, 43: 155–167.
- Suñer, M. (1992). Subject clitics in the Northern Italian vernaculars and the matching hypothesis. *Natural Language and Linguistic Theory*, 10: 641–672.
- Svenonius, P. (2002). *Subjects, expletives, and the EPP*. New York: Oxford University Press.
- Travis, L. (1984). *Parameters and Effects of Word Order Variation*. Ph.D. dissertation, MIT.
- Vikner, S. (1995). *Verb movement and expletive subjects in the Germanic Languages*. Oxford: Oxford University Press.
- Zanuttini, R. (1997). *Negation and Clausal Structure. A Comparative Study of Romance Languages*. Oxford: Oxford University Press.



## Appendix: Task Instructions

### **Instructions: Evaluation des phrases**

*Vous allez entendre des phrases en français et déterminer si elles vous paraissent acceptables. Par phrase acceptable, on veut dire qu'un locuteur natif du français pourrait l'utiliser dans une conversation naturelle entre personnes qui se connaissent bien. Par inacceptable on veut dire que la phrase vous paraîtrait bizarre si un ami l'utilisait dans une conversation avec vous, ce qui vous mènerait à penser que le français n'est pas sa langue maternelle.*

*On ne s'intéresse ni aux points de style ou de clarté, ni aux règles de grammaire que vous avez pu apprendre à l'école. Ce sont vos premières intuitions en tant que locuteur français qui nous intéressent, donc concentrez-vous seulement sur ce qui vous paraît plus ou moins naturel en conversation. De plus, on s'intéresse particulièrement à ce qui est naturel dans le **style 'informel' du français** -- la langue que vous utilisez en conversation avec vos amis proches ou votre famille, pas le style formel ou professionnel qu'on utilise quand on parle à son patron ou à ses profs.*

*Chaque phrase que vous allez juger sera précédée d'une question qui la situe dans un contexte particulier. Par exemple, vous pourriez entendre la question «Quel est votre film préféré ?», puis la réponse «Le film que j'aime le mieux est 8 Femmes». C'est la réponse, **la deuxième phrase**, que vous devez juger.*

*Utilisez une échelle de 1 (totalement inacceptable) à 7 (tout à fait acceptable) pour juger les phrases.*

*Essayez de ne pas vous attarder sur aucune phrase et de réfléchir à votre évaluation; indiquez votre toute première réaction pour chaque phrase.*

*Ce sondage n'est pas un test de mémoire ou un piège quelconque. N'essayez pas de mémoriser les jugements que vous donnez, ce sondage n'est pas un test secret d'intelligence! On est des linguistes, et on étudie la langue. On veut simplement savoir si vous trouvez acceptable ou non les phrases présentées pour que nous, on puisse mieux comprendre la langue française d'aujourd'hui.*

[ Translation:

*You will hear some sentences in French and decide whether they seem acceptable to you. By an acceptable sentence, we mean that a native speaker of French could use it in the natural conversation between people who know each other well. By unacceptable, we mean that the phrase would be strange to you if a friend used it in conversation with you, and it might lead you to think that French was not their native language.*

*We are not interested in points of style or clarity, nor in the rules of grammar that you may have learned in school. It's your immediate intuitions as a French speaker that interest us, so concentrate only on what would seem more or less natural to you in conversation. In addition, we are particularly interested in what is natural in 'informal' French—the language that you use in conversation with your close friends, not the formal or professional style that you might use when talking to your boss or professors.*

*Each phrase that you judge will be preceded by a question that situates it in a particular context. For example, you might hear the question “What’s your favorite movie?”, then the response “My favorite movie is ‘Huit Femmes’”. It's the response, the second sentence, that you must judge.*

*Use a scale from 1 (totally unacceptable) to 7 (perfectly acceptable) to judge the sentences.*

*Try not to take too long on any one sentence or reflect on your judgment; give your immediate reaction for each sentence.*

*This task is not a test of your memory or any other trick. Don't try to memorize the judgments that you've given, this task isn't a secret intelligence test! We are linguists, and we study language. We simply want to know whether you find the sentences we present to you acceptable or not, so that we can better understand the French language today. ]*